

AN ANALYSIS OF A CLASS OF VARIATIONAL MULTISCALE METHODS BASED ON SUBSPACE DECOMPOSITION

RALF KORNUBER, DANIEL PETERSEIM, AND HARRY YSERENTANT

ABSTRACT. Numerical homogenization tries to approximate the solutions of elliptic partial differential equations with strongly oscillating coefficients by functions from modified finite element spaces. We present a class of such methods that are closely related to the methods that have recently been proposed by Målqvist and Peterseim [Math. Comp. 83, 2014]. Like these methods, the new methods do not make explicit or implicit use of a scale separation. Their comparatively simple analysis is based on the theory of additive Schwarz or subspace decomposition methods.

1. INTRODUCTION

Numerical homogenization aims at a modification of standard finite element discretizations that preserve the accuracy known from smooth coefficient functions to the case of highly oscillatory coefficient functions. A method of this kind, which utilizes no separation of scales at all and is founded on a comprehensive convergence theory, has recently been proposed by Målqvist and one of the authors in [7]. The central idea of this paper is to assign to the vertices of the finite elements new basis functions that span the modified finite element space and reflect the multiscale structure of the problem under consideration. In the basic version of the method proposed there, the new basis functions are the standard piecewise linear hat functions minus their orthogonal projection onto a space of rapidly oscillating functions, orthogonal with respect to the symmetric, coercive bilinear form associated with the boundary value problem. They possess a global support but decay exponentially from one shell of elements surrounding the assigned vertex to the next. It is therefore possible to replace them by local counterparts without sacrificing the accuracy. The support of these localized basis functions consists of a fixed number of shells of elements surrounding the associated node. The number of these shells increases logarithmically with increasing accuracy, that is, decreasing gridsize.

In the paper here, we present a considerably simplified analysis of a class of closely related methods that is based on the theory of iterative methods, more precisely of additive Schwarz or subspace decomposition methods [9], [10], and essentially utilizes a reinterpretation [8] of the method proposed in [7] in terms of variational multiscale methods. Arguments used in the present paper are inspired by the work of Demko [3] on the inverses of band matrices and have been used, for example, in the proof of the H^1 -stability of the L_2 -orthogonal projection onto finite element spaces [1]. Our proof works on a more abstract level than that in [7] and is less centered around the single finite element basis functions. Implicitly, it

2010 *Mathematics Subject Classification.* 65N12, 65N30, 65N55.

This research was supported by Deutsche Forschungsgemeinschaft through SFB 1114.

compares their projections onto the mentioned space of rapidly oscillating functions with their iteratively calculated approximations and shows in this way that these decay exponentially with the distance to the assigned nodes. It rests, as that of all related results, upon a local version of the Poincaré inequality and thus depends on the local contrast ratio. A different class of methods [2], [5] not suffering from such restrictions is based on hierarchical matrices [4].

2. THE EQUATION AND THE BASIC APPROXIMATION OF ITS SOLUTIONS

The model problem considered in this paper is a standard second order differential equation in weak form with homogeneous Dirichlet boundary conditions on a polygonal domain Ω in $d = 2$ or 3 space dimensions. Its solution space is the Sobolev space $H_0^1(\Omega)$ and the associated bilinear form reads

$$(2.1) \quad a(u, v) = \int_{\Omega} \nabla u \cdot A \nabla v \, dx.$$

The matrix A is a function of the spatial variable x with measurable entries and is assumed to be symmetric positive definite. We assume for simplicity that

$$(2.2) \quad \delta |\eta|^2 \leq \eta \cdot A(x) \eta \leq M |\eta|^2$$

holds for all $\eta \in \mathbb{R}^d$ and almost all $x \in \Omega$, where $|\eta|$ denotes the euclidean norm of η and δ and M are positive constants. In the same way as in [6] it is possible to replace this condition by a at least in the quantitative sense weaker local condition on the contrast ratio. The condition (2.2) guarantees that the bilinear form (2.1) is an inner product on $H_0^1(\Omega)$. It induces the energy norm $\|\cdot\|$, which is equivalent to the original norm on this space. The boundary value problem

$$(2.3) \quad a(u, v) = f^*(v), \quad v \in H_0^1(\Omega),$$

possesses by the Lax-Milgram theorem under the condition (2.2) for all bounded linear functionals f^* on $H_0^1(\Omega)$ a unique solution $u \in H_0^1(\Omega)$.

We cover the domain Ω with a triangulation \mathcal{T} , consisting of triangles in two and of tetrahedrons in three space dimensions. We assume that the elements in \mathcal{T} are shape regular but do not require that \mathcal{T} is quasiuniform. Associated with \mathcal{T} is the conforming, piecewise linear finite element subspace \mathcal{S} of $H_0^1(\Omega)$. A key ingredient of the methods discussed here is a bounded local linear projection operator

$$(2.4) \quad \Pi : H_0^1(\Omega) \rightarrow \mathcal{S} : u \rightarrow \Pi u$$

like that defined as follows. At first, the given function $u \in H_0^1(\Omega)$ is locally, on the single elements $t \in \mathcal{T}$, approximated by its L_2 -orthogonal projection onto the space of linear functions, without regard to the continuity across the boundaries of the elements. In a second step, the values of these approximants at a vertex in the interior of the domain are replaced by their arithmetic mean and their values at a vertex on the boundary by the value zero. These values characterize the projection Πu of u onto \mathcal{S} . For the functions u in the solution space $H_0^1(\Omega)$, then the estimates

$$(2.5) \quad |\Pi u|_1 \leq c_1 |u|_1, \quad \|h^{-1}(u - \Pi u)\|_0 \leq c_2 |u|_1$$

hold, where $\|\cdot\|_0$ denotes the L_2 -norm, $|\cdot|_1$ the H^1 -seminorm, and h is an element-wise constant function whose value on the interior of a given element is its diameter. The first condition means that the projection operator (2.4) is stable with respect to the H^1 -norm and therefore also with respect to the energy norm. The second condition is an approximation property. Because such constructions are common

in finite element theory, we omit the proof here and refer only to [6] for a related result. The constants c_1 and c_2 depend, as with any other reasonable choice of Π , only on the shape regularity of the finite elements, but not on their size.

The kernel $\mathcal{V} = \ker \Pi$ of Π is a closed subspace of $H_0^1(\Omega)$ and therefore itself a Hilbert space. We can therefore introduce the a -orthogonal projection operator C from $H_0^1(\Omega)$ onto the kernel of Π and moreover the finite dimensional subspace

$$(2.6) \quad \mathcal{W} = \{v - Cv \mid v \in \mathcal{S}\}$$

of the a -orthogonal complement of the kernel of Π . The dimension of \mathcal{W} and of the finite element space \mathcal{S} coincide as $v \in \mathcal{S}$ can be recovered from $v - Cv$ via

$$(2.7) \quad v = \Pi(v - Cv).$$

In [8], the equation (2.3) is discretized using \mathcal{W} both as trial and test space.

Lemma 2.1 (Peterseim [8], cp. also Målqvist and Peterseim [7]). *The equation*

$$(2.8) \quad a(w, \chi) = f^*(\chi), \quad \chi \in \mathcal{W},$$

possesses a unique solution $w \in \mathcal{W}$, namely the projection

$$(2.9) \quad w = \Pi u - C\Pi u$$

of the exact solution $u \in H_0^1(\Omega)$ of equation (2.3) onto the space \mathcal{W} . The error

$$(2.10) \quad u - w = Cu$$

is the a -orthogonal projection Cu of the solution u onto the kernel of Π .

Proof. As Π is a projection operator, $u - \Pi u$ is contained in the kernel of Π . The difference of the exact solution u and the function (2.9) can therefore be written as

$$u - w = (u - \Pi u) - C(u - \Pi u) + Cu = Cu.$$

Because the functions in \mathcal{W} are a -orthogonal to the functions in the range of C ,

$$a(u - w, \chi) = 0, \quad \chi \in \mathcal{W},$$

follows. That is, this w is the unique solution of the equation (2.8). \square

The energy norm of Cu can easily be estimated for right hand sides

$$(2.11) \quad f^*(v) = \int_{\Omega} f v \, dx.$$

One obtains in this way the following, rather surprising error estimate.

Theorem 2.2 (Målqvist and Peterseim [7]). *For right hand sides of the form (2.11), the approximate solution (2.9) satisfies the energy norm error estimate*

$$(2.12) \quad \|u - w\| \leq c \|hf\|_0,$$

where the constant c depends only on the constants c_2 from (2.5) and δ from (2.2).

Proof. The proof is based on the representation (2.10) of the error as a -orthogonal projection Cu of the solution u onto the kernel of Π and starts from the identity

$$\|Cu\|^2 = a(u, Cu - \Pi Cu) = (f, Cu - \Pi Cu),$$

from which one obtains the estimate

$$\|Cu\|^2 \leq \|hf\|_0 \|h^{-1}(Cu - \Pi Cu)\|_0.$$

The second factor on the right hand side is estimated with help of the error estimate from (2.5), and the H^1 -seminorm of Cu then with (2.2) by its energy norm. \square

Remarkably, neither the smoothness of the solution u nor the regularity properties of the equation enter into this error estimate. The size of the error bound is determined by the local behavior of the right hand side f . Using the same type of arguments, one can prove the weighted L_2 -error estimate

$$(2.13) \quad \|h^{-1}(u - w)\|_0 \leq c^2 \|hf\|_0,$$

with the same constant c as in (2.12). That is, for uniform grids the convergence order in the L_2 -norm doubles in comparison to that in the energy norm.

3. LOCALIZATION

Let x_1, x_2, \dots, x_n be the vertices of the elements in the triangulation \mathcal{T} and let $\varphi_1, \varphi_2, \dots, \varphi_n$ be the piecewise linear hat functions assigned to these nodes. The φ_i assigned to the nodes in the interior of the domain Ω form then a basis of the finite element space \mathcal{S} , and the corresponding functions $\varphi_i - C\varphi_i$ a basis of the trial space (2.6). It has been shown in [7] that these basis functions decay exponentially with the distance to the assigned nodes and can therefore be replaced by localized counterparts. We deviate here from the argumentation there and utilize the theory of iterative methods to prove a result of similar kind. Let ω_i be the union of the finite elements with vertex x_i and let

$$(3.1) \quad \mathcal{V}_i = \{v - \Pi v \mid v \in H_0^1(\omega_i)\}.$$

The functions in \mathcal{V}_i vanish outside a small neighborhood of the vertex x_i , depending on the choice of Π . For the exemplary operator mentioned earlier, this neighborhood consists of the two shells of elements surrounding x_i . The \mathcal{V}_i are closed subspaces of the kernel \mathcal{V} of Π . This can be seen as follows. Let the $v_k \in \mathcal{V}_i$ converge to the function $v \in \mathcal{V}$. As the v_k are piecewise linear outside ω_i , the same holds for v . Thus there exists a function φ in the finite element space \mathcal{S} such that $v - \varphi \in H_0^1(\omega_i)$. Because $\Pi v = 0$ and $\Pi\varphi = \varphi$, therefore

$$(3.2) \quad v = (v - \varphi) - \Pi(v - \varphi) \in \mathcal{V}_i.$$

Let P_i be the a -orthogonal projection from $H_0^1(\Omega)$ to \mathcal{V}_i , defined via the equation

$$(3.3) \quad a(P_i v, v_i) = a(v, v_i), \quad v_i \in \mathcal{V}_i.$$

Introducing the operator

$$(3.4) \quad T = P_1 + P_2 + \dots + P_n,$$

the approximation spaces replacing \mathcal{W} are built up with the help of the bounded linear operators F_ν from $H_0^1(\Omega)$ to \mathcal{V} that are, starting from $F_0 u = 0$, defined via

$$(3.5) \quad F_{\nu+1} u = F_\nu u + T(u - F_\nu u).$$

The correction $T(u - F_\nu u)$ is the sum of its components $d_i = P_i(u - F_\nu u)$ in the subspaces \mathcal{V}_i of \mathcal{V} , the solutions $d_i \in \mathcal{V}_i$ of the local equations

$$(3.6) \quad a(d_i, v_i) = a(u, v_i) - a(F_\nu u, v_i), \quad v_i \in \mathcal{V}_i.$$

The new trial and test spaces are the spaces \mathcal{W}_ℓ spanned by the functions

$$(3.7) \quad \varphi_i - F_\nu \varphi_i, \quad \nu = 0, 1, \dots, \ell,$$

attached to the nodes x_i in the interior of the domain Ω . In contrast to their counterparts $\varphi_i - C\varphi_i$ spanning the original space \mathcal{W} they have a local support, which expands layer by layer with the number ν of iterations.

To study the approximation properties of these spaces \mathcal{W}_ℓ , we consider optimally or almost optimally chosen fixed linear combinations

$$(3.8) \quad C_\ell = \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} F_\nu, \quad \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} = 1,$$

of the operators F_ν as approximations of the a -orthogonal projection C . These operators C_ℓ serve here solely as a tool and do not need to be explicitly accessible. Our analysis is based on the theory of additive Schwarz or subspace decomposition methods, here applied to an equation in the kernel \mathcal{V} , and starts from the following

Lemma 3.1. *For all $v \in \mathcal{V}$, there is a with respect to the energy norm stable decomposition $v = v_1 + \dots + v_n$ of v into functions v_i in the spaces \mathcal{V}_i , such that*

$$(3.9) \quad \sum_{i=1}^n \|v_i\|^2 \leq K_1 \|v\|^2$$

holds, where the constant K_1 depends only on the constants c_1 and c_2 from (2.5), on the shape regularity of the finite elements, and on the contrast ratio M/δ . Moreover, there is a constant K_2 such that

$$(3.10) \quad \|v\|^2 \leq K_2 \sum_{i=1}^n \|v_i\|^2$$

holds for all such decompositions of v into functions v_i in the subspaces \mathcal{V}_i of the kernel. This constant depends only on the shape regularity of the finite elements.

Proof. The upper estimate (3.10) is rather trivial because K_2 can be bounded in terms of the maximum number of the parts v_i that do not vanish on a given element. We use that the φ_i form a partition of unity and prove that (3.9) holds for the decomposition of a function v in the kernel \mathcal{V} of Π into the parts

$$v_i = \varphi_i v - \Pi(\varphi_i v), \quad i = 1, \dots, n,$$

in \mathcal{V}_i . It suffices to prove that this decomposition is H^1 -stable. By the H^1 -stability of the projection Π and the shape regularity of the finite elements one obtains

$$\sum_{i=1}^n |v_i|_1^2 \lesssim \sum_{i=1}^n |\varphi_i v|_1^2 \lesssim |v|_1^2 + \|h^{-1}v\|_0^2.$$

Using once more that $\Pi v = 0$, the second term on the right hand side can, by the approximation property from (2.5), be estimated as

$$\|h^{-1}v\|_0 = \|h^{-1}(v - \Pi v)\|_0 \lesssim |v|_1.$$

Since the energy norm can be estimated by the H^1 -seminorm and vice versa, the estimate (3.9) follows and the proof is complete. \square

With the help of the estimates (3.9) and (3.10) one can show that

$$(3.11) \quad 1/K_1 a(v, v) \leq a(Tv, v) \leq K_2 a(v, v)$$

holds for the functions v in the kernel \mathcal{V} of Π . The spectrum of T , seen as an operator from \mathcal{V} to itself, is therefore a compact subset of the interval with the endpoints $1/K_1$ and K_2 . Because $I - F_\nu = (I - T)^\nu$ and $F_\nu C = F_\nu$,

$$(3.12) \quad C - C_\ell = \left\{ \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} (I - T)^\nu \right\} C.$$

Using the spectral mapping theorem and the fact that the norm of a bounded, symmetric operator from a Hilbert space to itself is equal to its spectral radius, one gets therefore similarly to the finite dimensional case the following error estimate.

Lemma 3.2. *If the weights $\alpha_{\ell\nu}$ are optimally chosen, the estimate*

$$(3.13) \quad \|Cu - C_\ell u\| \leq \frac{2q^\ell}{1 + q^{2\ell}} \|Cu\|$$

holds for all $u \in H_0^1(\Omega)$, where the convergence rate

$$(3.14) \quad q = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

is determined by the condition number $\kappa \leq K_1 K_2$ of the operator (3.4) seen as bounded, symmetric operator from the subspace \mathcal{V} of $H_0^1(\Omega)$ to itself.

The distance of $C_\ell u$ to Cu can thus be estimated in terms of the two constants K_1 and K_2 and the norm of Cu and tends exponentially in the number ℓ of iterations to zero. The lemma is basically a reformulation of a standard result from the theory of subspace decomposition methods [9], [10], where Cu is here the solution of the equation and the $C_\ell u$ are its iteratively generated approximations. We refer to [6] for the missing details of the here only sketched proof.

Theorem 3.3. *Let w and w_ℓ be the best approximations of the solution u of the original equation (2.3) in \mathcal{W} and \mathcal{W}_ℓ with respect to the energy norm. Then*

$$(3.15) \quad \|u - w_\ell\| \leq \left(1 + \frac{2q^\ell}{1 + q^{2\ell}}\right) \|u - w\| + \frac{2q^\ell}{1 + q^{2\ell}} \|u - \Pi u\|.$$

Proof. Because w_ℓ is the best approximation of u by a function in \mathcal{W}_ℓ , because the function $\Pi u - C_\ell \Pi u$ is contained in this space, and because $w = \Pi u - C\Pi u$,

$$\|u - w_\ell\| \leq \|u - (\Pi u - C_\ell \Pi u)\| = \|(u - w) - (C\Pi u - C_\ell \Pi u)\|.$$

The distance of $C\Pi u$ and $C_\ell \Pi u$ can according to Lemma 3.2 be estimated by the energy norm of $C\Pi u$. As $C\Pi u = (u - w) - (u - \Pi u)$, this leads to (3.15). \square

We conclude that logarithmically many iteration steps ν or even less, depending on the behavior of the energy norm of $u - \Pi u$, suffice to reach the same level of accuracy as with the original space \mathcal{W} based on the exact projection C . The best approximation w_ℓ of the solution u in the space \mathcal{W}_ℓ can itself again be calculated iteratively by an additive or multiplicative Schwarz or subspace decomposition method based on the splitting of \mathcal{W}_ℓ into the finite element space \mathcal{S} , that provides for the global exchange of information, and local subspaces like

$$(3.16) \quad \mathcal{W}_{\ell,i} = \text{span} \{ \varphi_i - F_\nu \varphi_i \mid \nu = 0, \dots, \ell \},$$

bearing the information on the fine scale structure of the solution.

The proof of Theorem 3.3 shows that one can replace the $\ell + 1$ functions (3.7) attached to the nodes x_i in the interior of Ω by a single, fixed linear combination

$$(3.17) \quad \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} (\varphi_i - F_\nu \varphi_i), \quad \sum_{\nu=0}^{\ell} \alpha_{\ell\nu} = 1,$$

without sacrificing the error bound (3.15). This considerably reduces the size of the spaces \mathcal{W}_ℓ . The problem is that the optimum coefficients $\alpha_{\ell\nu}$ depend on the

entire spectrum of the operator T from the kernel \mathcal{V} to itself. In a much simplified variant, approximations C_ℓ of the projection C are determined via a recursion

$$(3.18) \quad C_{\ell+1}u = C_\ell u + \omega T(u - C_\ell u),$$

where $C_0 u = 0$ is set and $\omega \geq 1/K_2$ is a damping parameter whose optimal value depends again on the end points of the spectrum. Because in this case

$$(3.19) \quad C - C_\ell = (I - \omega T)^\ell C,$$

convergence is guaranteed at least for $\omega < 2/K_2$, and for $\omega = 1/K_2$ in particular. The functions $\varphi_i - C_\ell \varphi_i$ spanning the new spaces \mathcal{W}_ℓ can be calculated in the same way as the functions $\varphi_i - F_\nu \varphi_i$. The convergence rate degrades, however, with this version from (3.14) to a value not better than

$$(3.20) \quad q = \frac{\kappa - 1}{\kappa + 1},$$

and in the extreme case to $q = 1 - 1/(K_1 K_2)$ if $\omega = 1/K_2$ is chosen. This means that possibly a larger number ℓ of iterations and layers, respectively, is needed.

4. DISCRETIZATION

The infinite dimensional subspaces \mathcal{V}_i of the kernel of the projection Π have to be replaced by discrete counterparts to obtain a computationally feasible method. We start from a potentially very strong, uniform or nonuniform refinement \mathcal{T}' of the triangulation \mathcal{T} , bridging the scales and resolving the oscillations of the coefficient functions, and a finite element space $\mathcal{S}' \subseteq H_0^1(\Omega)$ that consists of the continuous functions whose restrictions to the elements in \mathcal{T}' are linear. The whole theory then literally transfers to the present situation replacing only the continuous solution space $H_0^1(\Omega)$ and its subspaces by their discrete counterparts. The only modification concerns the construction of the stable decomposition of the functions v in the kernel of Π into a sum of functions in the corresponding local subspaces

$$(4.1) \quad \mathcal{V}_i = \{v - \Pi v \mid v \in \mathcal{S}' \cap H_0^1(\omega_i)\}.$$

To construct such a decomposition, we use the interpolation operator $\mathcal{I} : C(\bar{\Omega}) \rightarrow \mathcal{S}'$ that interpolates at the nodes of usual kind and reproduces the functions in \mathcal{S}' . As the operator \mathcal{I} is linear and the φ_i form a partition of unity, we can decompose the functions $v \in \mathcal{S}'$ in the kernel of Π into the sum of the functions

$$(4.2) \quad v_i = \mathcal{I}(\varphi_i v) - \Pi(\mathcal{I}(\varphi_i v)),$$

in the modified subspaces \mathcal{V}_i . The stability of this decomposition in the sense of (3.9) can be deduced in the same way as the stability of the decomposition in the proof of Lemma 3.1 since for the functions $v \in \mathcal{S}'$ in the kernel of Π an estimate

$$(4.3) \quad |\mathcal{I}(\varphi_i v)|_1 \leq c |\varphi_i v|_1$$

holds, which is shown separately for the single elements $t \in \mathcal{T}'$ using that the restrictions of such functions $\varphi_i v$ to these elements are second order polynomials.

REFERENCES

1. R. E. Bank and H. Yserentant, *On the H^1 -stability of the L_2 -projection onto finite element spaces*, Numer. Math. **126** (2014), 361–381.
2. M. Bebendorf, *Low-rank approximation of elliptic boundary value problems with high-contrast coefficients*, SIAM J. Math. Anal. **48** (2016), 932–949.
3. S. Demko, *Inverses of band matrices and local convergence of spline projection*, SIAM J. Numer. Anal. **14** (1977), 616–619.
4. W. Hackbusch, *Hierarchical Matrices: Algorithms and Analysis*, Springer Series in Computational Mathematics, vol. 49, Springer, 2015.
5. W. Hackbusch and F. Drechsler, *Partial evaluation of the discrete solution of elliptic boundary value problems*, Comput. Visual. Sci. **15** (2012), 227–245.
6. R. Kornhuber and H. Yserentant, *Numerical homogenization of elliptic multiscale problem by subspace decomposition*, Multiscale Model. Simul. **14** (2016), 1017–1036.
7. A. Målqvist and D. Peterseim, *Localization of elliptic multiscale problems*, Math. Comp. **83** (2014), 2583–2603.
8. D. Peterseim, *Variational multiscale stabilization and the exponential decay of fine-scale correctors*, Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations (G. R. Barrenechea, F. Brezzi, A. Cangiani, and E. H. Georgoulis, eds.), Lecture Notes in Computational Science and Engineering, vol. 114, Springer, 2016, pp. 341–367.
9. J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), 581–613.
10. H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta Numerica **2** (1993), 285–326.

INSTITUT FÜR MATHEMATIK, FREIE UNIVERSITÄT BERLIN, 14195 BERLIN, GERMANY
E-mail address: kornhuber@math.fu-berlin.de

INSTITUT FÜR NUMERISCHE SIMULATION, UNIVERSITÄT BONN, 53115 BONN, GERMANY
E-mail address: peterseim@ins.uni-bonn.de

INSTITUT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT BERLIN, 10623 BERLIN, GERMANY
E-mail address: yserentant@math.tu-berlin.de